Molecular Evolution and Host Adaptation of *Bordetella* spp.: Phylogenetic Analysis Using Multilocus Enzyme Electrophoresis and Typing with Three Insertion Sequences

ANNEKE VAN DER ZEE. 18 FRITS MOOI. 1 JAN VAN EMBDEN. 1 AND JAMES MUSSER 2

Laboratory of Infectious Disease Research, National Institute of Public Health and Environment, 3720 BA Bilthoven,
The Netherlands, and Section of Molecular Pathobiology, Department of Pathology,
Baylor College of Medicine, Houston, Texas 770302

Received 4 March 1997/Accepted 30 August 1997

A total of 188 Bordetella strains were characterized by the electrophoretic mobilities of 15 metabolic enzymes and the distribution and variation in positions and copy numbers of three insertion sequences (IS). The presence or absence of IS elements within certain lineages was congruent with estimates of overall genetic relationships as revealed by multilocus enzyme electrophoresis. Bordetella pertussis and ovine B. parapertussis each formed separate clusters, while human B. parapertussis was most closely related to IS1001-containing B. bronchiseptica isolates. The results of the analysis provide support for the hypothesis that the population structure of Bordetella is predominantly clonal, with relatively little effective horizontal gene flow. Only a few examples of putative recombinational exchange of an IS element were detected. Based on the results of this study, we tried to reconstruct the evolutionary history of different host-adapted lineages.

The genus *Bordetella* consists of the following species; *B*. pertussis, B. parapertussis, B. bronchiseptica, and B. avium. These four species are pathogens of the respiratory tract and have adapted to a variety of vertebrate hosts. B. avium causes respiratory disease in fowl, while the other three species are associated with disease in mammals. B. pertussis is an obligate human pathogen causing whooping cough or pertussis. B. parapertussis generally causes a milder form of human pertussis and has also been isolated from sheep with chronic nonprogressive pneumonia (9). B. bronchiseptica causes kennel cough in dogs and atrophic rhinitis in pigs and is associated with disease in a broad variety of other mammals. B. bronchiseptica is only rarely associated with disease in humans but can cause infections in immunocompromised patients (35). Recently, two additional Bordetella species have been identified and designated B. holmesii and B. hinzi (8, 34). Both of these species were isolated from the blood of immunocompromised human patients. In contrast to the other four Bordetella species, these organisms were not associated with antecedent respiratory disease (8, 34).

The *Bordetella* species can be distinguished based on biochemical properties, growth characteristics, and motility. In addition, within one *Bordetella* species, subgroups can be differentiated. For example, *B. parapertussis* ovine and human isolates can be differentiated by phenotypical properties, multilocus enzyme electrophoresis (MEE) (28), and insertion sequence (IS)-generated DNA polymorphism (32). *B. bronchiseptica* strains can be differentiated by the presence or absence of IS1001 (31).

Several studies have shown that *B. pertussis*, human *B. parapertussis*, and *B. bronchiseptica* are very closely related genetically (2, 14, 15, 19). By MEE, the genetic diversity among *B. pertussis*, *B. parapertussis* (isolated from humans), and *B. bronchiseptica* was found to be very limited (24). On the basis of overall genomic character and probable phylogenetic relation-

ships, it was proposed that *B. pertussis* and human isolates of *B. parapertussis* should be considered clones of *B. bronchiseptica* that adapted to the human host relatively recently (24). By MEE analysis, strains are assigned to particular electrophoretic types (ETs). The number of polymorphisms is roughly proportional to the time of divergence from a common ancestor. An alternative approach in determining the genetic relationships of bacterial strains is by analyzing strains with respect to the distribution and variation in positions and copy numbers of ISs (4, 16). Within different ETs, a finer resolution can often be obtained with IS elements, because of the faster rate of change. In *Bordetella*, three IS elements, IS481 (22), IS1001 (31), and IS1002 (32), have been identified and shown to be polymorphic in copy number and chromosomal position (32, 33).

Several questions remain unanswered about the phylogeny of *Bordetella* species. Did *B. pertussis* and *B. parapertussis* have a common ancestor that adapted to the human host or did they evolve from *B. bronchiseptica* strains associated with different nonhuman mammalian hosts? Is ovine *B. parapertussis* most closely related to human *B. parapertussis*, and did it evolve from the human *B. parapertussis* population by adaptation to another host or vice versa?

In this study we attempted to establish the fine-structure molecular evolutionary relationships between *Bordetella* strains in relation to host adaptation by determining the distribution of three different IS elements within different multilocus genotypes of *Bordetella*.

MATERIALS AND METHODS

Bacterial strains. A total of 188 *Bordetella* strains were analyzed by MEE and investigated for the presence of either IS1001, IS1002, or IS481. These strains included 18 *B. pertussis* isolates, 10 human *B. parapertussis* isolates, 16 ovine *B. parapertussis* isolates, and 144 *B. bronchiseptica* strains isolated from various hosts as indicated in Table 1. The strains studied represent the breadth of lineages of the species, but not all individual lineages were studied. Strains with ETs 2, 5, 7, 9, 10–13, 15, and 17–21 described previously (25) were no longer available and, hence, could not be investigated for the presence of IS elements. Strains with the aforementioned ETs are not listed in Table 1. However, to be thorough, we chose to include these 14 ETs of *B. bronchiseptica* in the dendrogram in Fig. 2.

^{*} Corresponding author. Present address: Laboratory of Molecular Microbiology, St. Elisabeth Hospital, P.O. Box 747, 5000 AS Tilburg, The Netherlands. E-mail: Lab.med.microbiol@inter.NL.net.

TABLE 1. Bordetella strains analyzed in this study

Species	Host	IS present	ET	Strain(s) ^a	n	Geographic origin(s)
B. bronchiseptica	Pig	IS <i>1001</i>	1	586, 603, 604, 605, 606, 653, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 805, 806, 846, 847, 848, 849, 850, 853, 963, 976, 978, 979, 980	34	The Netherlands, United Kingdom, Sweden, United States, Russia, Germany, South Africa, Australia, Switzerland, Ireland
	Pig	<u></u> b	1	553, 548	2	United Kingdom, The Netherlands
	Pig	IS1001	3	654, 676, 677, 688, 804	5	The Netherlands, United States, Australia
	Pig	IS1001	27	545	1	USSR
	Rabbit	IS <i>1001</i>	1	589, 704, 823, 826, 828, 833, 834, 836, 838, 972, 981	11	Finland, Switzerland, Denmark, The Netherlands, South Africa, United States
	Rabbit	_	3	784	1	The Netherlands
	Rabbit	_	16	730, 831, 835, 837, 971	5	Denmark, Switzerland, The Netherlands
	Rabbit	_	30	705	1	United States
	Rabbit	_	31	671	1	United States
	Rabbit	_	32	824, 832	2	Switzerland
	Cat	IS1001	1	635, 745, 969	3	United States, Denmark, The Netherlands
	Cat	-	16	629, 630, 631, 723, 733, 968, 970	7	United States, Denmark, The Netherlands
	Cat		32	782	1	The Netherlands
	Dog	IS <i>1001</i>	1	685, 686, 785, 965, 966	4	United States, The Netherlands
	-	131001	4	595	1	· · · · · · · · · · · · · · · · · · ·
	Dog	_				United States United States The Notherlands Switzer
	Dog	_	6	590, 592, 593, 594, 596, 597, 599, 600, 601, 602, 786, 826, 827, 843, 851, 852	16	United States, The Netherlands, Switzer- land
	Dog	_	8	732, 748, 788, 839, 967	5	Denmark, The Netherlands, Switzerland
	Dog	_	14	591	1	United States
	Dog		16	749, 750, 783, 787, 892, 977	6	The Netherlands, South Africa, Denmark
	Guinea pig	IS1001	1	674, 762, 854	3	Germany, Ireland
	Guinea pig	IS481	1	678	1	Australia
	Guinea pig	IS1001	3	673	1	Germany
	Guinea pig	_	3	670	1	United States
	Guinea pig		16	627, 665, 666, 667, 668, 669	6	United States
	Koala	IS1001	1	680, 681, 698, 700	4	Australia
	Koala	IS1002	1	679	1	Australia
	Horse	_	16	624, 628, 632, 633, 898, 983	6	United States, Germany
	Horse	IS481	34	731	1	Denmark
	Horse	_	34	982	1	United Kingdom
	Cow	IS1001	1	759	1	Ireland
	Seal	_	1	959, 960	2	Denmark
	Mouse	IS1001	1	973	1	The Netherlands
	Unknown	IS1001	1	974, 975	2	The Netherlands, South Africa
	Rat	_	16	625	1	United States
	Leopard	_	16	626	1	United States
	Turkey	_	16	901	1	Germany
	Turkey	_	29	707	1	United States
	Monkey	_	33	902	1	Germany
	Human	_	14	675	1	Germany
B. parapertussis	Sheep	IS1001	22	NZ929	1	New Zealand
	Sheep	IS1001	23	SC2209, SC2080, SC2225, SC2235	4	Scotland
	Sheep	IS <i>1001</i>	24	SC6, SC7, SC8, SC9, SC10, SC1991, SC2222, NZ928	8	Scotland, New Zealand
	Sheep	IS1001	25	NZ585, NZ927	2	New Zealand
	Sheep	IS <i>1001</i>	26	SC11	1	Scotland
	Human	IS1001/IS1002	28	B24, B265, B266, B270, B271, B279, B280, B281, 531, 718	10	The Netherlands, Finland, United States, Germany, New Zealand
B. pertussis	Human	IS481/IS1002	35	B6	1	The Netherlands
1	Human	IS481/IS1002	36	B3, B5, B12, B17, B18, B28, B389, B390, B391, B396, B397, B398, 814, 884, 885	15	The Netherlands, United States
	Human	IS481/IS1002	37	B44 (Tohama)	1	Japan
	Human	IS481/IS1002	38	B89 (18-323)	1	United States

^a Strains from collection of J. M. Musser are designated by numbers, and those from the National Institute of Public Health and Environment are preceded by letters. ^b—, no IS element found.

Growth of bacteria and electrophoresis of enzymes. Bacteria were harvested from Bordet-Gengou agar plates and subsequently grown in 500 ml of Stainer-Scholte medium at 37°C on an orbital shaker. Lysates were prepared for protein gel electrophoresis as described previously (25). Lysates were electrophoresed on starch gels and stained for the following 15 enzymes; NAD-dependent malate dehydrogenase, fumarase, indophenol oxidase, hydroxybu-

tyrate dehydrogenase, alkaline phosphatase (ALP), esterase, phosphoglucose isomerase, isocitrate dehydrogenase (IDH), phosphoglucomutase (PGM), adenylate kinase (ADK), leucine aminopeptidase (LAP), glutamate dehydrogenase, leucylalanine peptidase, catalase, and glutamic oxaloacetic transaminase (GOT). Different buffers and pHs were used in electrophoresis as described previously (25). Each isolate was characterized by the combination

TABLE 2. Allele profiles of 15 enzyme loci of Bordetella

ET	Reference	n	Electrophoretic mobilities of each enzyme ^b											Species				
EI	isolate ^a	n	MDH	FUM	IPO	HBD	ALP	EST	PGI	IDH	PGM	ADK	LAP	GLD	PE1	CAT	GOT	species
1	586	69	1	2	1	2	2	2	1	1	3	1	3	1	3	1	2	B. bronchiseptica
2	ET2		1	2	1	2	1	2	1	1	3	1	3	1	3	1	2	B. bronchiseptica
3	654	8	1	2	1	2	2	2	1	1	4	1	3	1	3	1	2	B. bronchiseptica
4	595	1	1	1	1	2	2	2	2	1	3	1	3	1	3	1	2	B. bronchiseptica
5	ET5		1	2	1	2	2	2	2	1	3	1	3	1	3	1	2	B. bronchiseptica
6	590	16	1	1	1	2	2	1	1	1	3	1	3	1	3	1	2	B. bronchiseptica
7	ET7		1	2	1	2	2	1	1	1	3	1	3	1	3	1	2	B. bronchiseptica
8	732	5	1	1	1	2	2	2	1	1	3	1	3	1	3	1	2	B. bronchiseptica
9	ET9		1	1	1	2	2	4	1	1	3	1	3	1	3	1	2	B. bronchiseptica
10	ET10		1	1	1	2	2	1	1	1	3	1	2	1	3	1	2	B. bronchiseptica
11	ET11		1	1	1	2	2	1	1	1	3	1	3	2	3	1	2	B. bronchiseptica
12	ET12		1	1	1	2	2	2	2	1	4	1	2	1	3	1	2	B. bronchiseptica
13	ET13		1	2	1	2	1	2	1	1	1	1	2	1	1	1	2	B. bronchiseptica
14	591	2	1	2	1	2	1	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
15	ET15		1	2	1	2	1	2	1	1	3	1	2	1	3	1	2	B. bronchiseptica
16	730	33	1	2	1	2	2	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
17	ET17		1	2	1	2	4	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
18	ET18		1	2	1	2	5	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
19	ET19		1	2	1	1	2	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
20	ET20		1	2	1	2	2	3	1	1	3	1	2	1	1	1	2	B. bronchiseptica
21	ET21		1	1	1	1	3	2	1	1	2	2	3	2	2	1	1	B. bronchiseptica
22	NZ929	1	1	2	1	2	1	2	1	0	3	1	3	1	3	1	1	B. parapertussis (sheep)
23	SC2209	4	1	2	1	2	1	2	1	3	3	1	3	1	3	1	2	B. parapertussis (sheep)
24	SC6	8	1	2	1	2	2	2	1	3	3	1	3	1	3	1	2	B. parapertussis (sheep)
25	NZ585	2	1	2	1	2	2	2	1	3	3	1	3	1	3	1	1	B. parapertussis (sheep)
26	SC11	1	1	2	1	2	2	2	1	1	3	1	3	1	3	1	1	B. parapertussis (sheep)
27	545	1	1	2	1	2	2	2	1	1	3	1	4	1	3	1	2	B. bronchiseptica
28	B24	10	1	2	1	2	2	2	1	1	3	3	3	1	3	1	2	B. parapertussis (human)
29	707	1	1	1	1	2	2	1	1	1	4	1	3	1	3	1	2	B. bronchiseptica
30	705	1	1	1	1	2	2	2	1	1	3	1	2	1	3	1	2	B. bronchiseptica
31	671	1	1	1	1	2	2	2	1	1	3	1	1	1	3	1	2	B. bronchiseptica
32	824	3	1	1	1	2	2	2	1	1	3	1	2	1	1	1	2	B. bronchiseptica
33	902	1	1	1	1	2	1	2	1	1	3	1	2	1	3	1	2	B. bronchiseptica
34	731	2	2	2	1	2	2	2	1	1	3	1	2	1	3	1	2	B. bronchiseptica
35	B6	1	1	2	1	2	2	2	1	1	3	3	1	1	1	1	2	B. pertussis
36	B3	15	1	2	1	2	2	2	1	1	3	3	1	1	1	1	3	B. pertussis
37	B44	1	1	2	1	2	2	2	1	0	5	3	1	1	1	1	2	B. pertussis
38	B89	1	1	2	1	2	2	2	1	1	5	1	3	1	1	1	2	B. pertussis
50	D 03	1	1	4	1	_	_	4	1	1	5	1	5	1	1	1	4	D. periussis

^a Reference isolates designated "ET and number" refer to historical isolates that are no longer available.

of alleles at the 15 enzyme loci and with relation to previously established ETs (25), which were used as controls.

Determination of the presence of IS elements. Part of the strain collection had already been blotted and hybridized to IS-specific probes (32, 33). These strains included all *B. pertussis* (33) and *B. parapertussis* isolates (32) used in this study and 50 of the *B. bronchiseptica* strains, 41 of which contained IS1001 (32). All additional strains were investigated for the presence of IS1001, IS1002, and IS481 by PCR as described previously (32).

Statistical analysis. The genetic diversity at a locus (h) among ETs or isolates was calculated as follows: $h = (1 - \Sigma x_i^2)(n/n - 1)$, where x_i is the frequency of the ith allele and n is the number of ETs or isolates. Mean genetic diversity per locus is the arithmetic average of h values over all loci. Genetic distance between pairs of ETs was calculated as the proportion of loci at which different alleles were represented, and a dendrogram of ETs was constructed from a matrix of genetic distances by the average linkage method (30).

Analysis of multilocus linkage disequilibrium was carried out as described by Maynard Smith et al. (21) based on the method of Brown et al. (7).

The number of IS-specific hybridizing bands was determined by Southern blotting. Because the intensities of hybridizing bands were similar, single bands were assumed to contain one copy of an IS element. Bands of the same size were assumed to be identical. To measure the genetic variation between banding patterns created by restriction fragment length polymorphism (RFLP), we devised the following equation. The genetic diversity based on RFLP analysis was calculated as $d = 1 - [2x_{ab}/(x_a + x_b)]$, where x_{ab} is the number of bands of identical size between two isolates a and b, and x_a and x_b are the number of bands in isolates a and b, respectively. The mean genetic divergence between clones was calculated by comparing all pairs of RFLP types.

RESULTS

Genetic diversity among ETs and isolates of Bordetella. For the 18 isolates of B. pertussis, 4 of 15 enzyme loci were polymorphic (PGM, ADK, LAP, and GOT), and four ETs were identified (Tables 2 and 3). The mean (± standard deviation) genetic diversity per locus among ETs was 0.178 ± 0.079, and that among isolates was 0.056 ± 0.024 . The human B. parapertussis isolates were the same ET. Among 16 ovine B. parapertussis isolates, 3 of 15 enzymes were polymorphic (ALP, IDH, and GOT) (Tables 2 and 3). Five different ETs were identified. The mean genetic diversity per locus was 0.127 ± 0.068 among ETs, and that among isolates was 0.076 ± 0.043 . Among B. bronchiseptica strains, 14 ETs were identified, 7 of which were assigned to B. bronchiseptica isolates analyzed in a previous study (25). Seven new ETs were found. So as not to underestimate the genetic diversity among B. bronchiseptica isolates, we included the ETs that were found before (25). When all ETs of B. bronchiseptica were included, the mean genetic diversity per locus among a total of 28 ETs was 0.232 ± 0.054 . The overall mean genetic diversity of 38 ETs of Bordetella was 0.252 ± 0.049 .

^b Allele denotations 0 to 5 refer to the different electrophoretic mobilities of each enzyme. MDH, NAD-dependent malate dehydrogenase; FUM, fumarase; IPO, indophenol oxidase; HBD, hydroxybutyrate dehydrogenase; EST, esterase; PGI, phosphoglucose isomerase; CAT, catalase; PE1, leucylalanine peptidase.

TABLE 3.	Multilocus linkage disequilibrium in ETs of	
	Bordetella (sub)species	

	,	, -		
(Sub)species (host)	Cluster ^a	No. of ETs	D^b	$I_A{}^c$
All Bordetella species		38	0.252	0.464 ± 0.22
B. parapertussis and B. bron- chiseptica (mainly pig)	A	10	0.142	-0.123 ± 0.42
B. parapertussis (sheep)	AB	5	0.127	0.300 ± 0.52
B. parapertussis (human) and B. bronchiseptica (mainly pig)	AC	5	0.107	-0.309 ± 0.57
B. bronchiseptica (dog)	D	12	0.156	0.210 ± 0.38
B. bronchiseptica (broad host)	E	11	0.147	-0.153 ± 0.42
B. pertussis (human)	F	4	0.178	0.684 ± 0.64

- ^a Clusters B and C are joined within cluster A.
- ^b D, mean genetic distance per locus between ETs.

Distribution of IS1001, IS1002, and IS481 among Bordetella strains. All strains were investigated for the presence of IS elements by PCR. IS1001 was present in all B. parapertussis strains and in 70 of 144 B. bronchiseptica strains (32) (Table 1). The number of copies of IS1001 among ovine B. parapertussis isolates ranged from 23 to 29, and four RFLP types were identified (Fig. 1 and Table 4). The strains of these 4 RFLP types had 11 IS1001-containing SphI restriction fragments. Among human isolates of B. parapertussis, IS1001 was present in 20 to 21 copies, and three different RFLP types were identified among which 15 bands of identical sizes were present (Fig. 1). IS1001 was present in 70 of 144 B. bronchiseptica strains. Among B. bronchiseptica strains, the copy number of IS1001 ranged from 1 to 7. Only one band was common to 41 of 70 strains, and 21 different RFLP types were found. Thus, based both on the variation in IS1001 copy number and on IS1001-based RFLP analysis, genotypic diversity was most limited among human B. parapertussis isolates and most varied among isolates of B. bronchiseptica. Nine copies of IS1002 were present in human isolates of B. parapertussis, and the RFLP patterns of all 10 isolates were identical. IS1002 was not found in ovine B. parapertussis isolates and occurred in only 1 of the 144 B. bronchiseptica isolates (a koala bear isolate), in which only a single copy was present. B. pertussis strains contained IS1002, and the copy number ranged from 4 to 8 per cell (Fig. 1). In 100 B. pertussis strains, 36 different IS1002-based RFLP types were identified (33).

IS481 is present in all *B. pertussis* isolates with an extremely high copy number which is estimated to be between 80 and 100 (12, 22). IS481 is absent from *B. parapertussis* but was found in 2 of 144 *B. bronchiseptica* strains (from a guinea pig and a horse), with one and two copies, respectively. In summary, genotypic diversity, as estimated by RFLP analysis, was most limited among human isolates of *B. parapertussis*, with higher values for isolates of *B. pertussis*, ovine *B. parapertussis*, and *B. bronchiseptica* (Fig. 1 and Table 4).

Reconstruction of a phylogenetic tree. A dendrogram generated by the average linkage method presents estimates of the genetic relationships among the 38 multilocus genotypes (Fig. 2). The presence of IS481, IS1001, and IS1002 within ETs is indicated.

Surprisingly, all IS1001-containing Bordetella isolates were

clustered within one group of related ETs (cluster A), which is composed of two subgroups of ETs (clusters B and C). Cluster B contains all ovine B. parapertussis isolates, and each ET within this cluster was associated with a single IS1001 RFLP pattern, except for strains of ET 24, among which three different RFLP types were found. Cluster C is comprised of ETs of B. bronchiseptica isolates (ETs 1 to 3 and $27\overline{)}$ and human B. parapertussis isolates (ET 28). Among a total of 78 B. bronchiseptica isolates assigned to cluster C, 70 contained IS1001. Two IS1001-negative B. bronchiseptica strains in this cluster had either IS1002 or IS481. ETs 3 and 27 contained isolates with IS1001-based RFLP types found in strains within ET 1. All human B. parapertussis isolates within ET 28 had nine copies of IS1002 (one RFLP type) in addition to IS1001 (three RFLP types). Clusters D and E and ET 21 comprised ETs of B. bronchiseptica isolates, and all isolates except one were negative for the presence of any of the three IS sequences. Only one B. bronchiseptica isolate of ET 34 (cluster E) had two copies of IS481. Cluster F comprised ETs of all B. pertussis isolates (ET 35 to 38). All B. pertussis isolates had a high number of copies of IS481 and four to eight copies of IS1002. Most isolates of B. pertussis were of ET 36 and showed similar, but not identical, IS1002-based RFLP patterns. B. pertussis of ETs 35 and 37 showed different RFLP patterns, while the observed genetic distance of B. pertussis strain 18-323 (ET 38) was reflected by its IS1002-based RFLP type (Fig. 1).

Although different options in ET clustering are possible by average linkage algorithm, especially when genetic distances are ≥ 10 , the composition of main clusters was not affected, and the ET of human *B. parapertussis* strains always clustered with that of IS1001-containing *B. bronchiseptica* isolates.

Host range of Bordetella in particular clusters or ETs. B. parapertussis strains within cluster B (Fig. 2) are strictly confined to sheep, while B. parapertussis of ET 28 (cluster C) were recovered only from humans. Strains within cluster F are highly host specific because B. pertussis is found strictly in humans. Host adaptation also occurs among B. bronchiseptica strains but is less strict. For example, all pig isolates of B. bronchiseptica were assigned to cluster C, but strains derived from other hosts (guinea pigs, cats, rabbits, and dogs) are found occasionally in this cluster. B. bronchiseptica strains of ET 6 and ET 8 (cluster D) were isolated only from dogs, while other ETs within the same cluster contained isolates from rabbits (ET 30, 31) and a turkey (ET 29). B. bronchiseptica strains within ET 16 (cluster E) showed a very broad host range and were isolated from eight different hosts. In contrast to the B. bronchiseptica strains within ET 1, no preference for one particular host was evident, as within ET 16, similar numbers of strains were isolated from different hosts (Table 1 and Fig. 2).

The population structure of the genus Bordetella. Linkage disequilibrium analysis of all Bordetella ETs shows that the nonrandom association of alleles differs significantly from zero, indicating that recombinational events are very rare in natural populations of Bordetella (Table 3). However, when clusters of related ETs are subjected to linkage disequilibrium analysis, the evidence for nonrandom association of alleles largely disappears. The observation of reduced I_A (index of association) values correlates with the host specificity of Bordetella isolates within separate clusters; it approaches linkage equilibrium (IA = 0) within clusters of strains with a narrow host range (clusters B, D, and F) and indicates a freely recombining population $(I_{\Delta} < 0)$ when strains have a broader host range (clusters C and E). Thus, the I_A of alleles at different loci within separate clusters of ETs suggests that recombination may occur between strains of related lineages but not between more distantly related ones (Table 3).

^c Standard error on the index according to Maynard Smith et al. (21).

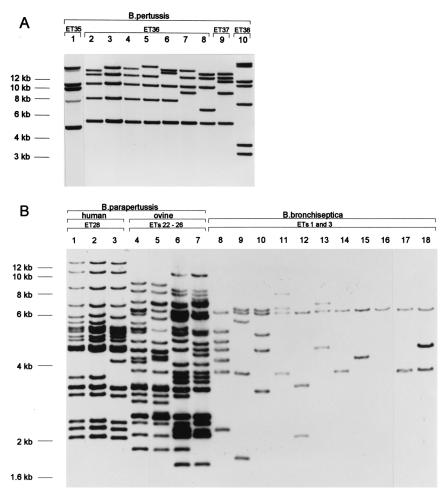


FIG. 1. IS1002-generated RFLP patterns among *B. pertussis* strains (A) and IS1001-generated RFLP patterns of human and ovine *B. parapertussis* and *B. bronchiseptica* strains (B). This figure represents a compilation of previously published (32, 33) and novel data. *B. pertussis* DNA was digested with *SmaI* and hybridized with IS1002. (A) *B. pertussis* strains: B6 (lane 1), B396 (lane 2), B397 (lane 3), B398 (lane 4), B389 (lane 5), B390 (lane 6), B410 (strain ND) (lane 7), B391 (lane 8), B44 (strain Nonama) (lane 9), B89 (18-323) (lane 10). *B. parapertussis* and *B. parapertussis* strains: B24 (lane 1), 531 (lane 2), B271 (lane 3), NZ585 (RFLP type 1) (lane 4), SC7 (RFLP type 2) (lane 5), SC10 (RFLP type 3) (lane 6), SC11 (RFLP type 4) (lane 7), *B. bronchiseptica* strains: 805 (lane 8), 969 (lane 9), 823 (lane 10), 799 (lane 11), 785 (lane 12), 973 (lane 13), 677 (lane 14), 759 (lane 15), 636 (lane 16), 676 (lane 17), 688 (lane 18). The ETs to which strains were assigned are indicated above the lanes. The sizes of marker DNA fragments are indicated on the left.

DISCUSSION

Overall genetic diversity and relationships among *Bordetella*. Our results, as determined by MEE, show that the overall level of genetic diversity of *Bordetella* is restricted compared to that of other bacterial pathogens (13). In general, populations

of pathogenic bacteria, such as *Escherichia coli* (20), *Haemophilus influenza* (27), and *Streptococcus pyogenes* (26), show substantial variation in genotypic diversity. The restricted genotypic diversity among *Bordetella* species suggests a very recent evolutionary origin. The high similarity of 23S rRNA gene

TABLE 4. Estimates of genetic diversity based on the distribution, abundance, and polymorphism generated by IS elements within *Bordetella* (sub)species

(Sub)species (host)	n	IS present	IS copy no.	No. of RFLP types/no. of isolates ^a	Genetic diversity ^b	
B. pertussis	18	IS481, IS1002	±100, 4-7	ND, 36/100	0.26	
B. parapertussis (human)	10	IS1001, IS1002	20–21, 9	3/20, 1/10	0.11	
B. parapertussis (sheep)	16	IS1001	23–29	4/22	0.39	
B. bronchiseptica (mainly pig)	70	IS1001	1–7	21/41	0.70	
B. bronchiseptica (guinea pig/horse)	2	IS481	1–2	2/2		
B. bronchiseptica (koala)	1	IS1002	1	1/1		
B. bronchiseptica (various hosts)	71					

^a Number of RFLP types per number of previously investigated isolates (32, 33). ND, not determined.

^b Calculated as $1 - (2x_{ab}/(x_a + x_b))$, where x_{ab} is the number of bands of identical size between two isolates a and b, and x_a and x_b are the number of bands in isolates a and b, respectively.

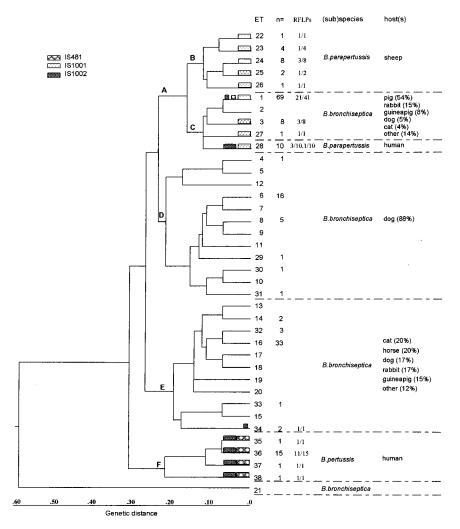


FIG. 2. Genetic relationships among *B. pertussis*, *B. parapertussis*, and *B. bronchiseptica* strains. The genetic distance between pairs of ETs was calculated as the proportion of loci at which different alleles were represented, and clustering of ETs was performed from a matrix of genetic distances by the average linkage method (17). A genetic distance of 0.07 between ETs corresponds to a single locus difference, and the largest (0.59) corresponds to differences at 10 of 15 enzyme loci. The presence of IS elements is indicated by boxes. Small boxes indicate that the IS was present in only one isolate of the ET. Large boxes are shown in ETs in which all or >91% and >75% (in ET 1 and ET 3, respectively) of isolates contained the given IS element. In all other ETs, none of the isolates were found to contain an IS element. The number of strains (n=) assigned to a particular ET is indicated. For a number of branches no *n* value is given, as these ETs refer to historical samples that are no longer available. The number of different RFLP types per number of investigated strains is indicated, based on IS1001 among *B. bronchiseptica* and ovine *B. parapertussis*, on IS1001 and IS1002 among human *B. parapertussis*, and on IS1002 among *B. pertussis*. Also mentioned is the determined *Bordetella* species and the host(s) from which the strains were isolated.

sequences (>99%) found previously also confirms the close relationship between members of the genus Bordetella (23). Based on the restricted amount of genetic diversity, it has been argued that Bordetella should be classified into subspecies rather than species (23, 24) and that *B. pertussis* and human *B.* parapertussis should be regarded as clones of B. bronchiseptica. The genetic diversity of Bordetella spp., as estimated by the distribution and variation in copy number of IS elements within different lineages, is congruent with the results determined by MEE. B. bronchiseptica isolates show the highest degree of genetic diversity and are subdivided into three distinct clusters of related ETs. Both MEE and the distribution of IS1001 show that B. parapertussis is more closely related to B. bronchiseptica in cluster C than other ETs of B. bronchiseptica are. This result reinforces the notion that B. parapertussis is just another lineage of B. bronchiseptica. Ovine and human B. parapertussis isolates are closely related but distinct and show a

remarkable difference in genetic diversity. Based both on MEE and IS typing, ovine *B. parapertussis* isolates show a higher degree of divergence than human isolates. *B. pertussis* isolates cluster separately both by their ETs and by the distributions of IS481 and IS1002 and are related relatively distantly to *B. parapertussis*.

The evolution of *B. pertussis*. *B. pertussis* and *B. parapertussis* may have derived from distinct clones of *B. bronchiseptica* because the ETs of *B. pertussis* and *B. parapertussis* are only distantly related, and both species contain different IS elements. The genetic diversity of *B. pertussis* was relatively large compared to that of *B. parapertussis*, a finding which has been demonstrated by pulsed-field gel electrophoresis of chromosomal restriction fragments (14). In theory, the more varied genetic diversity of *B. pertussis* compared to that of *B. parapertussis* might be due to a larger population, but since population size is measured by isolation frequency, estimates of incidence

are likely to be biased towards *B. pertussis*. Because clinical symptoms of *B. parapertussis* infections are usually less severe, infections often go unnoticed. Immunological data showed that 40 to 90% of the population has antibodies against *B. parapertussis* (6). Furthermore, intensive immunization programs with whole-cell pertussis vaccines may have limited the size of the *B. pertussis* population, whereas that of *B. parapertussis* may not have been affected. Thus, the sizes of the *B. pertussis* and *B. parapertussis* populations may be comparable.

Clinical symptoms of whooping cough were first described in the 16th century (18), and *B. pertussis* was isolated in 1904 (5). Although *B. parapertussis* is easier to grow, it was not isolated until decades later in 1937 (10). Because of the larger genetic diversity of *B. pertussis*, demonstrated both by MEE and IS1002-based RFLP analysis, and because of the relatively distant relationship with *B. parapertussis*, we conclude that they do not have a recent common ancestor and most likely evolved from different lineages. Furthermore, the association of *B. pertussis* with humans probably predates the association of *B. parapertussis* with humans.

The evolution of human *B. parapertussis*. Human *B. parapertussis* isolates belong to a single ET, show little variation in IS1001 and pulsed-field gel electrophoresis patterns, and no variation in IS1002-based RFLP patterns. These observations suggest that insufficient time has passed for this host species to diversify, and consequently it is likely that *B. parapertussis* became associated with disease in humans only very recently. Since this study was initiated, many more *B. parapertussis* strains from worldwide locations have been subjected to RFLP analysis, confirming the unusual genetic homogenity among isolates of this species.

Computer analysis by average linkage method clustered the ETs of both human and ovine B. parapertussis with that of IS1001-containing B. bronchiseptica. RFLP analysis of human and ovine isolates in a previous study suggested that ovine and human B. parapertussis evolved independently from B. bronchiseptica, based both on the similarity of IS1001-based hybridization patterns among strains of one host species and the dissimilarity of hybridization patterns between the different host species and on the sequence homology of IS1001 (32). Since the ET of human B. parapertussis isolates always clusters with B. bronchiseptica ETs 1 and 27, the derivation of human B. parapertussis from a B. bronchiseptica pig strain was confirmed by MEE analysis. Although no genes of ovine B. parapertussis have been sequenced, comparisons of fimbriae (29), adenylate cyclase/hemolysin (3), pertactin (17), and pertussis toxin genes (2, 19) confirm both the close genetic relationship between B. bronchiseptica and human B. parapertussis and the relatively distant relationship with B. pertussis. The observed genetic distance between human B. parapertussis and B. bronchiseptica, dictated by differences in the mobilities of two enzymes as determined by MEE, suggests that there may be a missing link in the ancestry of human B. parapertussis. The very recent emergence of a highly successful clone of B. parapertussis may present an alternative possibility in the evolution of this human pathogen and may explain the lack of significant divergence.

The evolution of ovine *B. parapertussis*. Among ovine *B. parapertussis* isolates, five separate ETs were found, three each among isolates from Scotland and New Zealand. Thus, ovine *B. parapertussis* probably evolved from *B. bronchiseptica* earlier than human *B. parapertussis*. We can only speculate whether sheep from Scotland carried *B. parapertussis* to New Zealand. If so, New Zealand *B. parapertussis* isolates may since have diverged into ETs 22, 24, and 25, while Scotland isolates may have diverged into ETs 23, 24, and 26. Since ET 24 was assigned to isolates of both geographic locations, a strain of this

ET may have been ancestral to the New Zealand *B. parapertussis* population. The high degree of similarity between RFLP type 1 (New Zealand) and type 2 (Scotland) may support this ancestral connection.

Relative divergence of B. bronchiseptica. Geographically isolated populations may diverge differentially and result in overestimation of genetic diversity, which may be the case with B. bronchiseptica ETs. In B. bronchiseptica isolates from The Netherlands or the United States, which were sampled from a wide range of hosts, six and eight different ETs were found, respectively. Some ETs appear to be confined to certain geographic locations, e.g., ET 8, which is found in Europe but not in the United States. This observation was made in a previous study with a different collection of Bordetella strains (25). This geographic component is illustrated by the separation of the ET of the Japanese isolate among B. pertussis isolates. Although the RFLP type of strain Tohama (ET 37) is very similar to that of *B. pertussis* ET 36 isolates, analysis of RFLP types from a previous study groups this isolate with other strains from Japan (33). In summary, some B. bronchiseptica ETs may be geographically contained, while other ETs, such as ET 1 and ET 16, are distributed worldwide.

The reliability of the phylogenetic tree. Only in the absence of frequent recombination can a phylogenetic tree be constructed which may represent the evolutionary history of related bacteria. Linkage disequilibrium analysis based on multilocus genotypes showed that, although the overall population structure of *Bordetella* is clonal, recombination between strains of related genotype or particular hosts may be common. The possible occurrence of recombination was most prominent in *B. bronchiseptica* lineages. *B. bronchiseptica* strains, however, have been recovered from a wide variety of hosts and geographic locations. As recombination between spatially isolated populations is impossible, linkage equilibria may not be reliable in these cases. Evidence of recombination was weaker among ovine *B. parapertussis* derived from two separate locations and among *B. pertussis* isolates.

The distribution of IS elements in the genus *Bordetella* may be the outcome of both horizontal and vertical transmission. The route by which an IS element was acquired can be inferred from MEE analysis. Thus, the fact that a group of closely related strains contains a particular IS element can usually be taken as evidence that the element is an ancestral condition and was transferred by descent. The pattern of distribution of IS elements within clusters of related multilocus genotypes and the similarities between RFLP patterns of related strains within these clusters suggest that these three IS elements have been transmitted predominantly vertically, rather than horizontally. The results suggest that IS1001 was acquired by the common ancestor of the cluster that harbors B. parapertussis and certain B. bronchiseptica lineages. IS elements may also be lost by deletion events, and the absence of IS1001 from only a small fraction of B. bronchiseptica strains within this cluster may be indicative of such an event. In B. pertussis, vertical gene transfer is probably also responsible for the dissemination of IS481 and IS1002. In a number of cases, acquisition of an IS element is probably due to recent horizontal transfer, an assumption which may be supported by the lower copy numbers of IS481 (one and two) in B. bronchiseptica strains of ET 1 and ET 34, respectively, and one copy of IS1002 in a B. bronchiseptica strain of ET 1. IS1002 was further confined to human isolates of B. pertussis and B. parapertussis. Since the association of B. pertussis with humans probably predates that of B. parapertussis, IS1002 has most likely been horizontally transferred to a coinfecting B. parapertussis strain.

Thus, the population structure of Bordetella seems predom-

inantly clonal, with occasional episodes of horizontal transfer o recombinational exchange. Evolutionary trees for *Bordetella* have been reconstructed before, based on sequence comparison of toxin genes (2), MEE (24), or both (1). The phylogenetic tree reconstructed here is in congruence with the tree reconstructed by Altschul (1) with regard to clustering of all *B. pertussis* isolates and the clustering of *B. bronchiseptica* ETs 1 to 3 together with that of human *B. parapertussis* isolates.

Host diversity of Bordetella. In view of the restricted genetic diversity within *Bordetella*, the differences observed in the host adaptation of lineages are striking. B. pertussis and two distinct lineages of B. parapertussis are strictly associated with humans or sheep. A previous study showed that B. bronchiseptica of ET 1 and ET 6 were pig and dog specialist clones, respectively, whereas strains of ET 16 showed less host specificity (25). Although B. bronchiseptica belonging to cluster C are isolated principally from pigs, they are capable of colonizing other hosts. B. bronchiseptica strains of ET 16 were also isolated from a broad range of different hosts, but in contrast to ET 1 isolates, no preference was indicated for any particular host. Although the dominance of ET 1 reflects an extensive sampling of pigs, this ET was found in 30% of strains derived from other hosts. In addition to having a worldwide presence, B. bronchiseptica of ET 1 may therefore be considered a very successful clone.

Recently it was reported that in B. bronchiseptica strains associated with phylogenetic lineage, the expression of alcaligin (a component involved in iron acquisition) was repressed by the two-component bvg regulatory system, whereas in other strains, it was by independent (11). This example may demonstrate that fine-tuning a phenotypic characteristic could facilitate the occupation of a particular niche or host. It is conceivable that apart from by repression of alcaligin, many phenotypic traits may play a role in defining host specificity. In view of the probable recent association of B. parapertussis with humans, this species may have undergone a more abrupt adaptation, a fact which might be established by host-specific adhesins. As IS1002 is confined strictly to human pathogens of B. parapertussis and B. pertussis, there is an intriguing possibility that horizontal transfer of IS1002, perhaps with associated sequences, played a role in the adaptation of B. parapertussis to the human host or in the emergence of a highly successful clone.

ACKNOWLEDGMENTS

We thank W. Gaastra, J. Porter, and B. W. Manktelow for providing strains. Melicia Tjoa is acknowledged for practical assistance, and Vivek Kapur is acknowledged for technical recommendations.

This research was financially supported by a grant from the Dutch Organization for Scientific Research.

REFERENCES

- Altschul, S. F. Evolutionary trees for the genus *Bordetella*. 1989. J. Bacteriol. 171:1211–1213.
- Arico, B., R. Gross, J. Smida, and R. Rappuoli. 1987. Evolutionary relationships in the genus *Bordetella*. Mol. Microbiol. 1:301–308.
- Betsou, F., O. Sismeiro, A. Danchin, and N. Guiso. 1995. Cloning and sequencing of B. bronchiseptica adenylate cyclase hemolysin encoding gene: comparison with the B. pertussis gene. Gene 162:165–166.
- Biserčíc, M., and H. Ochman. 1993. The ancestry of insertion sequences common to Escherichia coli and Salmonella typhimurium. J. Bacteriol. 175: 7863–7868
- Bordet, J., and O. Gengou. 1906. Le microbe coqueluche. Ann. Inst. Pasteur 20:710-741.
- Borska, K., and M. Simkovicova. 1972. Studies on the circulation of *Bordetella pertussis* and *Bordetella parapertussis* in populations of children. J. Hyg. Epidemiol. Microbiol. Immunol. 16:159–172.

 Brown, A. H. D., M. W. Feldman, and E. Nevo. 1989. Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96:523–536.

- Cookson, B. T., P. Vandamme, L. C. Carlson, A. M. Larson, J. V. L. Sheffield, K. Kersters, and D. H. Spach. 1994. Bacteremia caused by a novel *Bordetella* species, "B. hinzii." J. Clin. Microbiol. 32:2569–2571.
- Cullinane, L. C., M. R. Alley, R. B. Marsall, and B. W. Manktelow. 1987. Bordetella parapertussis from lambs. N. Z. Vet. J. 35:175.
- Eldering, G., and P. Kendrick. 1937. Bacillus parapertussis: a species resembling both Bacillus pertussis and Bacillus bronchisepticus but identical with neither. J. Bacteriol. 35:561–572.
- Giardina, P. C., L.-A. Foster, J. M. Musser, B. J. Akerley, J. F. Miller, and D. W. Dyer. 1995. bvg repression of alcaligin synthesis in Bordetella bronchiseptica is associated with phylogenetic lineage. J. Bacteriol. 177:6058– 6062
- Glare, E. M., J. C. Paton, R. R. Premier, A. J. Lawrence, and I. T. Nisbet. 1990. Analysis of a repetitive DNA sequence from *Bordetella pertussis* and its application to the diagnosis of pertussis using the polymerase chain reaction. J. Clin. Microbiol. 28:1982–1987.
- Go, M. F., V. Kapur, D. Y. Graham, and J. M. Musser. 1996. Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. J. Bacteriol. 178:3934–3938.
- Khattak, M. N., and R. C. Matthews. 1993. Genetic relatedness of *Bordetella* species as determined by macrorestriction digests resolved by pulsed-field gel electrophoresis. Int. J. Syst. Bacteriol. 43:659–664.
- Kloos, W. E., N. Mohapatra, W. J. Dobrogosz, J. W. Ezell, and C. R. Manclark. 1981. Deoxyribonucleotide sequence relationships among *Bordetella* species. Int. J. Syst. Bacteriol. 31:173–176.
- Lawrence, J. G., D. E. Dykhuizen, R. F. Dubose, and D. L. Hartl. 1989. Phylogenetic analysis using insertion sequence fingerprinting in *Escherichia coli*. Mol. Biol. Evol. 6:1.
- Li, J., N. F. Fairweather, P. Novotny, G. Dougan, and I. G. Charles. 1992.
 Cloning, nucleotide sequence and heterologous expression of the protective outer membrane protein P.68 pertactin from *Bordetella bronchiseptica*.
 J. Gen. Microbiol. 138:1697–1705.
- Major, R. H. 1954. A history of medicine V-I, p. 423. Charles C. Thomas, Springfield, Ill.
- Marchitto, K. S., S. G. Smith, C. Locht, and J. M. Keith. 1987. Nucleotide sequence homology to pertussis toxin gene in *Bordetella bronchiseptica* and *Bordetella parapertussis*. Infect Immun. 55:497–501.
- Maslow, J. N., T. S. Whittam, C. E. Gilks, R. A. Wilson, M. E. Mulligan, K. S. Adams, and R. D. Arbeit. 1995. Clonal relationships among bloodstream isolates of *Escherichia coli*. Infect. Immun. 63:2409–2417.
- Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt. 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA 90:4384–4388.
- McLafferty, M. A., D. R. Harcus, and E. L. Hewlett. 1988. Nucleotide sequence and characterization of a repetitive DNA element from the genome of *Bordetella pertussis* with characteristics of an insertion sequence. J. Gen. Microbiol. 134:2297–2306.
- Muller, M., and A. Hildebrandt. 1993. Nucleotide sequences of the 23S rRNA genes from *Bordetella pertussis*, *B. parapertussis*, *B. bronchiseptica* and *B. avium*, and their implications for phylogenetic analysis. Nucleic Acids Res. 1142320
- Musser, J. M., E. L. Hewlett, M. S. Peppler, and R. K. Selander. 1986. Genetic diversity and relationships in populations of *Bordetella* spp. J. Bacteriol. 166:230–237.
- Musser, J. M., D. A. Bemis, H. Ishikawa, and R. K. Selander. 1987. Clonal diversity and host distribution in *Bordetella bronchiseptica*. J. Bacteriol. 169: 2793–2803
- 26. Musser, J. M., V. Kapur, S. Kanjilal, D. M. Musher, N. L. Barg, K. H. Johnston, P. M. Schlievert, J. Henrichsen, D. Gerlach, et al. 1993. Geographic and temporal distribution of two highly pathogenic clones of *Streptococcus pyogenes* expressing allelic variants of pyrogenic exotoxin A (scarlet fever toxin) J. Infect. Dis. 167:337–346.
- Musser, J. M., V. J. Rapp, and R. K. Selander. 1987. Clonal diversity in Haemophilus pleuropneumoniae. Infect. Immun. 55:1207–1215.
- Porter, J. F., K. Connor, and W. Donachie. 1994. Isolation and characterization of *Bordetella parapertussis* like bacteria from ovine lungs. Microbiology 140:255–261.
- 29. Savelkoul, P. H. M., D. P. G. de Kerf, R. J. Willems, F. R. Mooi, B. A. M. van der Zeijst, and W. Gaastra. 1996. Characterization of the fim2 and fim3 fimbrial subunit genes of Bordetella bronchiseptica: roles of Fim2 and Fim3 fimbriae in adhesion. Infect. Immun. 64:5098–5105.
- Selander, R. K., and B. R. Levin. 1980. Genetic diversity and structure in Escherichia coli populations. Science 210:545–547.
- van der Zee, A., C. Agterberg, M. van Agterveld, M. Peeters, and F. R. Mooi. 1993. Characterization of IS1001, an insertion sequence element of Bordetella parapertussis. J. Bacteriol. 175:141–147.
- 32. van der Zee, A., H. Groenendijk, M. Peeters, and F. R. Mooi. 1996. The differentiation of *Bordetella parapertussis* and *Bordetella bronchiseptica* from man and animals as determined by DNA polymorphism mediated by two different insertion sequence elements suggests their phylogenetic relation-

- ship. Int. J. Syst. Bacteriol. 46:640–647.
 33. van der Zee, A., S. Vernooij, M. Peeters, J. van Embden, and F. R. Mooi. 1996. Dynamics of the population structure of Bordetella pertussis as measured by IS1002 associated restriction fragment length polymorphism: comparison of pre- and postvaccination strains and global distribution. Microbiology **142**:3479–3485.
- 34. Weyant, R. S., D. G. Hollis, R. E. Weaver, M. F. M. Amin, A. G. Steigerwalt, S. P. O'Connor, A. M. Whitney, M. I. Daneshvar, C. W. Moss, and D. J. Brenner. 1995. Bordetella holmesii sp. nov., a new gram-negative species associated with septicemia. J. Clin. Microbiol. 33(1):1-7.
 35. Western P. E. H. J. A. Moselt. 1001. LIN. infection associated with
- 35. Woolfrey, B. F., and J. A. Moody. 1991. Human infections associated with Bordetella bronchiseptica. Clin. Microbiol. Rev. 4:243-255.